REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

2. REPORT DATE 5/27/93

3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE
An Automated Approach for Classifying Generic Terrian Features Using Digital Elevation Data

5. FUNDING NUMBERS

6. AUTHOR(S)
Linda H. Graff

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Topographic Engineering Center
ATTN:CETEC-LO
Fort Belvoir, VA 22060-5546

8. PERFORMING ORGANIZATION REPORT NUMBER
R-197

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release;
Distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The application of experimental SW to interface digital elevation data with knowledge representation and IP tools to automatically classify landforms. I will discuss classification theory, the terrain classification system I used in this work, the data and sites used to develop the classification method, the method itself, overall results, method limitations, and recommendations and finally conclusions.

14. SUBJECT TERMS
Terrain features, Digital Elevation Data, terrain Classification

15. NUMBER OF PAGES 13

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

An Automated Approach for Classifying Generic Terrain Features Using Digital Elevation Data

By Linda Graff

GOOD AFTERNOON

Viewgraph 1
The title of my talk this afternoon is "An Automated Approach to Classifying Generic Terrain Features Using Digital Elevation Data".

My work at the U.S. Army Topographic Engineering Center (TEC) involves the application of experimental SW to interface digital elevation data with knowledge representation and IP tools to automatically classify landforms

INTRODUCTION

Viewgraph 2
This afternoon I will discuss classification theory, the terrain classification system I used in this work, the data and sites used to develop the classification method, the method itself, overall results, method limitations and recommendations and finally conclusions.

CLASSIFICATION

Due to the subjective nature of many commonly recognized generic terrain terms, such as hill and plain, most terrain classification systems employ specific geomorphologic terms used by earth scientists, such as drumlin and alluvial fan. However, work in cognitive categorization theory shows that it is often easier to classify basic-level or generic objects than it is to provide a more specific classification.

Viewgraph 3
Rosch et al. (1976) and Rosch (1978) proposed a classification system that attempts to deal with objects that are difficult to categorize. Categories range from general or superordinate, at the uppermost level, to specific or subordinate, at the bottom level. The categories in this system often have best examples called "prototypes" and are defined in part by human perception and experience.

Central to the system proposed by Rosch et al. are basic-level categories and objects. Basic-level objects are basic in perception, function, communication and knowledge organization.

In general, objects that are atypical or differ greatly from the prototype are more difficult to classify than those that closely resemble the category prototype. This is especially important when dealing with terrain features which often differ greatly from the textbook examples and from each other.

Hoffman (1985) uses the concept of basic-level objects and categories in his treatment of "generic" topographic terms. He states that these terms, which include hill and plain, are rooted in perception, judgement and experience and are frequently used to communicate the perceptual form of terrain.

TERRAIN CLASSIFICATION

In an attempt to simplify the complex, and often subjective, problem of terrain classification, this study employs a two-class approach to terrain classification. The basic-level or generic terrain feature used in this study is a mount. The term "mount" is adapted from the U.S. Geological Survey's (USGS) proposed Digital Line Graph-Enhanced (DLG-E) definition as "a landmass that projects conspicuously above its surroundings."

Viewgraph 4
Mount represents such elevated terrain features as hills, mountains and ranges. The prototype mount is considered to be "well-defined." A well-defined mount is an isolated, elevated mass with a distinguishable boundary and a summit or peak. By using the term mount I don't have to deal with the ambiguity inherent in these other terms for instance - when does a hill become a mountain? or how many mountains does it take to make a range?

In addition to mounts, other features at the basic or generic level could include such things as plains and flats. However, this research collectively classifies all features other than mount as "non-mount".

In this categorization system we also see specific geomorphologic landforms that may be more familiar to terrain specialists and earth scientists at a more detailed level of categorization, while all terrain features can be grouped into a very general class referred to as landforms.

Information used in manual terrain classification is combined with automated techniques to classify mount/non-mount areas using digital elevation data as the sole data source.

Manual terrain classification often relies on isolating the feature of interest from the surrounding terrain, then measuring attributes associated with the isolated feature.

For instance, when using aerial photographs, the boundaries between landforms are often identified by breaks in slope which create apparent tonal and topographic changes. Information such as this can be incorporated into an automated terrain classification scheme.

Automated analysis techniques used in previous studies frequently employ general geomorphometry measures and/or critical points. General geomorphometry is the measurement of landform characteristics over a broad continuous surface and often relies on altitude and such derivatives as slope, aspect and curvature.

Studies using critical points extract information directly from the elevation data without computing derivatives or other measures. Critical points provide the maximum amount of information about a surface. Although called by different names, these points include: peaks, pits, ridges, ravines, passes, slopes, break points and flats.

DATA

The digital data are 7.5-minute-based digital elevation models or DEMs produced by the U.S. Geological Survey. These data have a 30-meter spacing between X and Y locations. Each DEM covers an area corresponding to a 1:24,000 scale topographic map.

Viewgraph 5
This viewgraph shows 10 sites that were used to develop the mount classification algorithm. These sites are scattered across the US, have differing geologic and geomorphic histories and vary in local relief from a low of 93m at Verona, WI to a high of 800 m.at Paradise Range, CA

Results of the automatic classification were compared to manual classifications for the selected DEMs using shaded relief, synthetic stereo images. The manual classifications were performed by scientists at TEC.

METHOD

Various data layers are created by extracting geomorphometric measures and critcal points. These layers are then combined to automatically classify mounts.

The automatic classification of mounts was performed in three phases. The first is a preprocessing phase, then application of the classification algorithm itself, followed by a postprocessing phase.

Preprocessing

Viewgraph 6
To preprocess the data, the elevation data are smoothed twice to minimize the effects of local highs and lows in the elevation values. The smoothed data are used to obtain new data layers that provided information similar to that used in manual terrain classification. This information includes percent slope and critical points.

Slope is used to provide a boundary between mount and non-mount areas, the value of which is related to the local relief of the site. Final analysis determined a boundary slope between mount and non-mount of 10 percent in areas with local relief greater than or equal 250m and 6 percent in lower relief areas.

Critical point analysis initially focused on peaks which are defined as a center elevation within a 3 x 3 neighborhood that is greater than all eight neighboring elevations. However, the identification criteria were too strict to extract many peaks. For this reason, ridge points were used as the critical points. A ridge point has a higher elevation than the elevation of its cardinal neighbors in an east-west or north-south direction.

Mount Classification Algorithm

Viewgraph 7
The goal of the classification algorithm is to have continuous areas represented as mounts with no gaps in the classification which is representative of a manual classification. The mount classification algorithm has four steps, each step uses the classification of the previous step as input:

1. Reclass Ridges. Assign a boundary slope between mount and non-mount areas based on local relief of the entire DEM so for areas with local relief >= 250m a boundary slope of 10% was used and in lower relief areas a boundary slope of 6% was used. Then classify all ridge point locations with a slope greater than the boundary slope as mounts and all other points as non-mounts.

2. Grow to Boundary. Examine a 3 x 3 window of mounts classified in Step 1 (Result 1) and percent slope to "grow" the mounts from the ridge points to the boundary slope as follows: if the center of the 3 x 3 window is a non-mount **and** its slope is greater than the boundary slope **and** any of its eight neighbors are mount then reclassify the center as mount. Continue to iterate thru the entire file until no more reclassifications are made.

3. Grow Uphill. Continue to "grow" the mounts classified in Step 2 (Result 2), by looking for uphill trends in the data using the original elevation data as follows: if a non-mount is encountered after an uphill trend is established (increasing elevation with mount values), then it is reclassified as mount. The entire area is processed first from left to right then from right to left.

4. Fill-in Flats. Fill in non-mount areas located within mounts by examining the mounts classified in Step 3 (Result 3) and a 3 x 3 window of original elevation data. This algorithm states that if all three neighbors to any side of a center non-mount value are classed as mount **and** the elevation of the central value is greater than its closest mount neighbor, then reclassify the center from non-mount to mount. Continue to iterate thru the entire file until no more reclassifications are made.

Postprocessing

Viewgraph 8
Early analysis of the manually derived boundaries with the automatic classifications showed that the automatic classification resulted in many small isolated clumps where isolated ridge points were located. To eliminate these clumps, all mounts less than 25 cells (roughly 150m x 150m) are sieved from the final result.

RESULTS

Preliminary results from my work show that the method I developed for mount classification was most successful in high-relief physiographic regions where the mounts were well-defined, i.e., where there was a sharp break in slope at the boundary between mount and non-mount areas. These were also the most easily distinguished in the manual classification.

The method was less successful in low-relief areas or in the areas were the geomorphic processes obscured the landforms like the glaciated areas of WI.

Next I would like to go through the results as applied to several different sites.

Viewgraph 9
The first site is Gettysburg, PA. The local relief for the Gettysburg DEM is 117m. This area has an overall rolling terrain with several low relief isolated hills and a long linear ridge in the north-west corner. The automatic classification for this area was the lowest of all the original ten sites chosen for study. One reason for this is that the linear ridge was totally missed in the automatic classification. This ridge was missed due to the critical point and boundary slope selection requirements.

Viewgraph 10
The second site is Post Oak Mountains, TX with a local relief of 140m. This site has a few isolated fat-topped mounts, in addition to a broad low-relief flat-topped mesa in the south-east quadrant of the DEM. This site was chosen for illustration because it illustrates the inability of the classification algorithm to fill in large flat-topped mounts.

Viewgraph 11
Mustang Mountains, AZ has a relatively high local relief of 641m. This site has a good mix of well-defined conical mounts with lower relief, poorly defined mounts in the southern quarter of the area. The automatic classification methods worked well on the well-defined mounts but was not as successful where the mount boundaries were less distinct.

Viewgraph 12
Since the initial results of this work using the original ten DEMs, the focus of the work has fallen on the SW US. This viewgraph shows the classification results as applied to the Drinkwater Lake, CA DEM. Drinkwater Lake, California has a high local relief of 861 meters and is located in the same geographic area as several of the original 10 DEMs used to develop the classification method. As was expected the classification algorithm performed well as you can see visually there is a very good correspondence between the two.

However, this viewgraph reflects several refinements made since the original data was processed. These include: 1) a boundary class of 12% instead of 10% was used due to the high local relief. This produced less merging of the mounts, but as could be expected it also resulted in more open areas within the mounts themselves. The second refinement was an image processing post process which helped deal with small non-mount areas located within the mounts. The areas were reclassed based on their size.

LIMITATIONS AND RECOMMENDATIONS

The results of the classification are highly dependent on three characteristics of the developed method.

Viewgraph 13
First, application of a universal approach that applies the same procedures to each DEM regardless of geographic location may not be desirable. If the methods of mount identification can be tailored to the area covered by the DEM, it is possible that the mount classification would be useful in more areas than suggested by this research. However, it is likely that many areas require a much deeper model of terrain classification than the two-class scheme used in this study.

Second, it appears that local neighborhood operators that examine one small window of information at a time can provide valuable information, as a first look at terrain. However, in many cases a 3 x 3 window, such as that used in this research, may be too small and restrictive for terrain features. Application of more regional operators, that examine the feature as a whole, may be required for accurate classification. This will become especially important if a more specific classification is desired.

Finally, the results of this research suggest that a boundary slope exists between mount and non-mount. As used in this study, this "critical value" is a function of the local relief of the area. Further investigation with additional DEMs is required to determine if there is a unique local relief cut off relating to a slope boundary between mount and non-mount areas or, if this too, is dependant on the area under investigation. As was the case with the Drinkwater Lake, CA DEM a higher boundary slope produced better results. Thus perhaps several boundary slopes are required.

Incorporation of high level knowledge-based procedures may help constrain and simplify the classification problem, thereby reducing the limitations of the current approach. These procedures can include regional knowledge about the area such as the physiographic region and climate, or local knowledge such as vegetation and landuse.

Relationships between knowledge such as this and terrain features have been studied by terrain analysts for many years. This knowledge can be used in a top-down approach to tailor the classification method used in a particular area to the features that are expected to be present.

Additional limitations to the current work may be imposed by the quality and resolution of the data source. Studies have shown that 7.5 minute-based DEMs are sufficient to extract large terrain features such as drainage basins, lakes and, in certain cases, mounts. However, it is insufficient for extracting detailed, local information such as gully shape. This type of information is frequently used by terrain analysts when performing manual feature classification.

It may be possible to extract information, such as gully shape, from higher quality and resolution data. However, until better data becomes readily available, it may be possible to extract similar detailed information, such as hydrography and vegetation, from other digital feature data sources.

As previously stated the automatic classification of generic terrain features such as mounts could also facilitate the automatic classification of more specific geomorphologic landforms. Evans (1987) states that a form must be isolated from its surroundings prior to a specific classification. By separating individual mounts from each other and from their surroundings it may be possible to apply additional measures to more specifically identify the feature.

Measures applied to individual mounts may include zonal elevation and slope statistics, as well as various surface characteristic measurements. For each of the mounts in this area measures have been taken for max and min Z, local relief, max and min slope and slope range, as well as the perimeter, area, volume and various measures of shape or roundness.

CONCLUSIONS

Viewgraph 14
In conclusion: The automatic classification of generic terrain features from digital data is a novel problem and initial results suggest that it is successful in high-relief physiographic regions.

Although the actual classification method has limitations, this research represents a first step toward automated classification of generic terrain features. This method of partitioning a space into coarse generic features addresses certain needs for symbolic, mid-level terrain characterization generated from a low-level characterization such as raw elevation data. The results may be improved by incorporating high-level knowledge to guide the classification process.

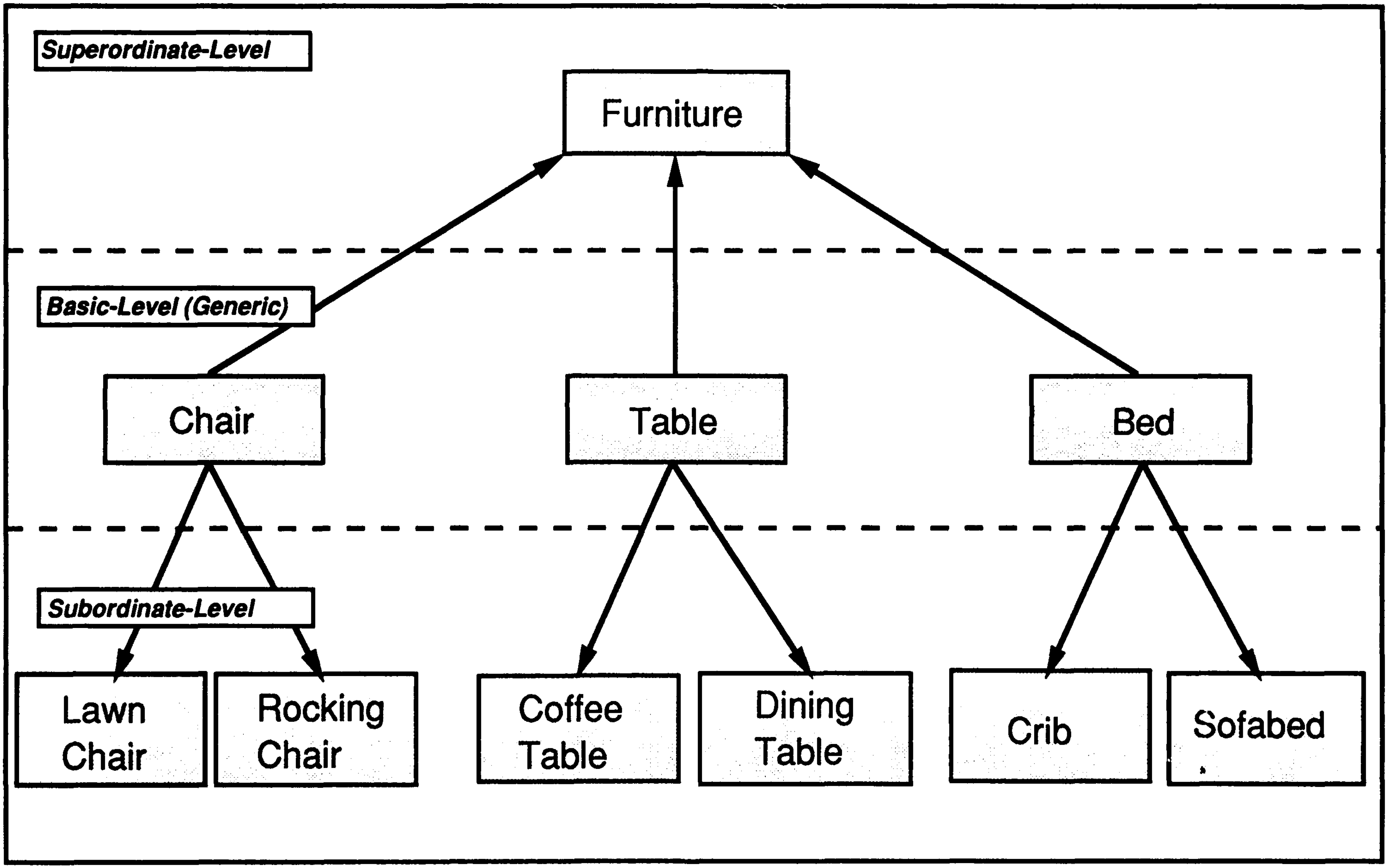
Also, by simplifying the classification problem and identifying a few generic terrain features, additional processing may lead to the identification of more specific geomorphologic landforms.

AN AUTOMATED APPROACH TO CLASSIFYING GENERIC TERRAIN FEATURES USING DIGITAL ELEVATION DATA

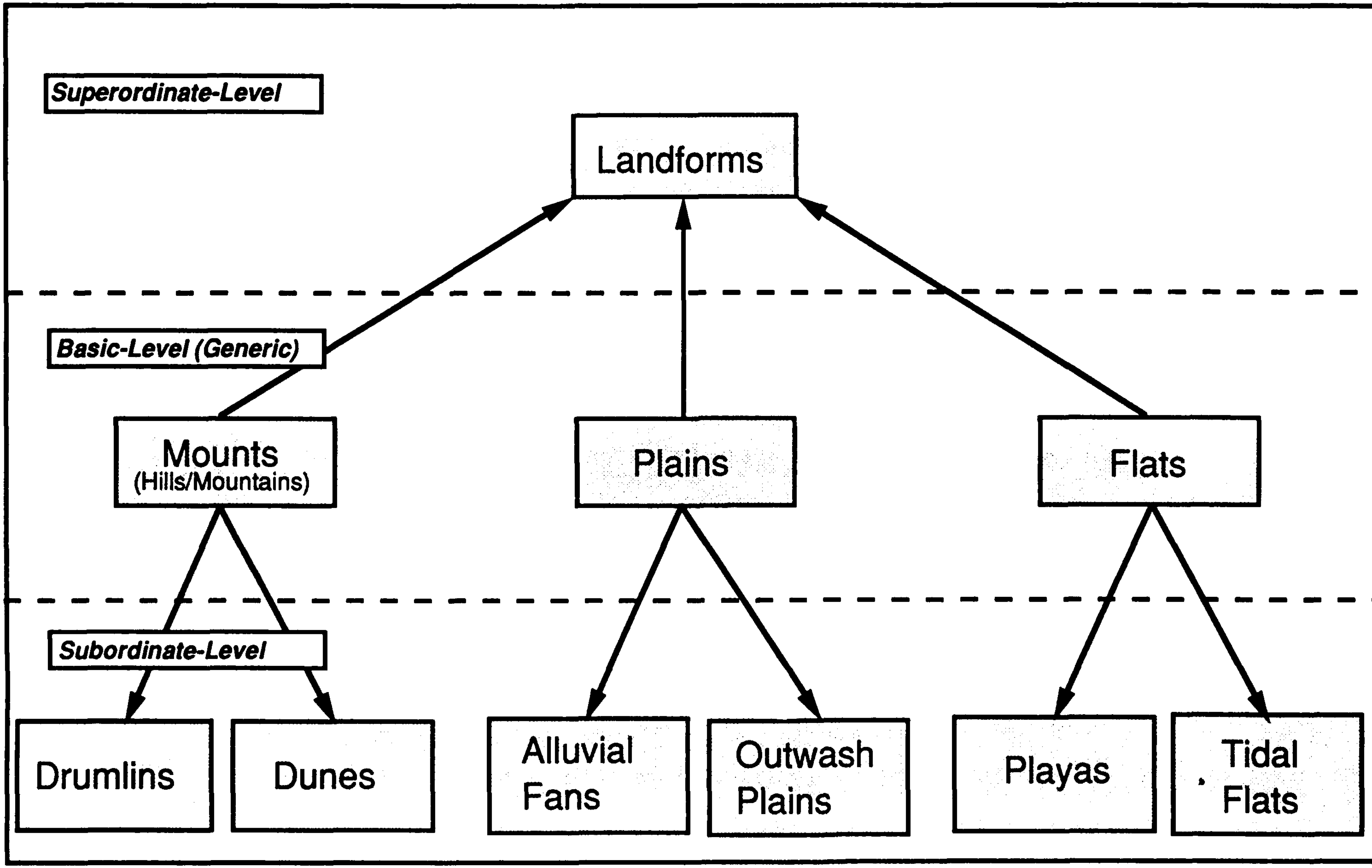
Linda Hoehn Graff
U.S. Army Topographic Engineering Center
Fort Belvoir, VA

- Introduction
- Classification Theory
- Terrain Classification
- Data and Sites Examined
- Method
- Results
- Recommendations
- Conclusions

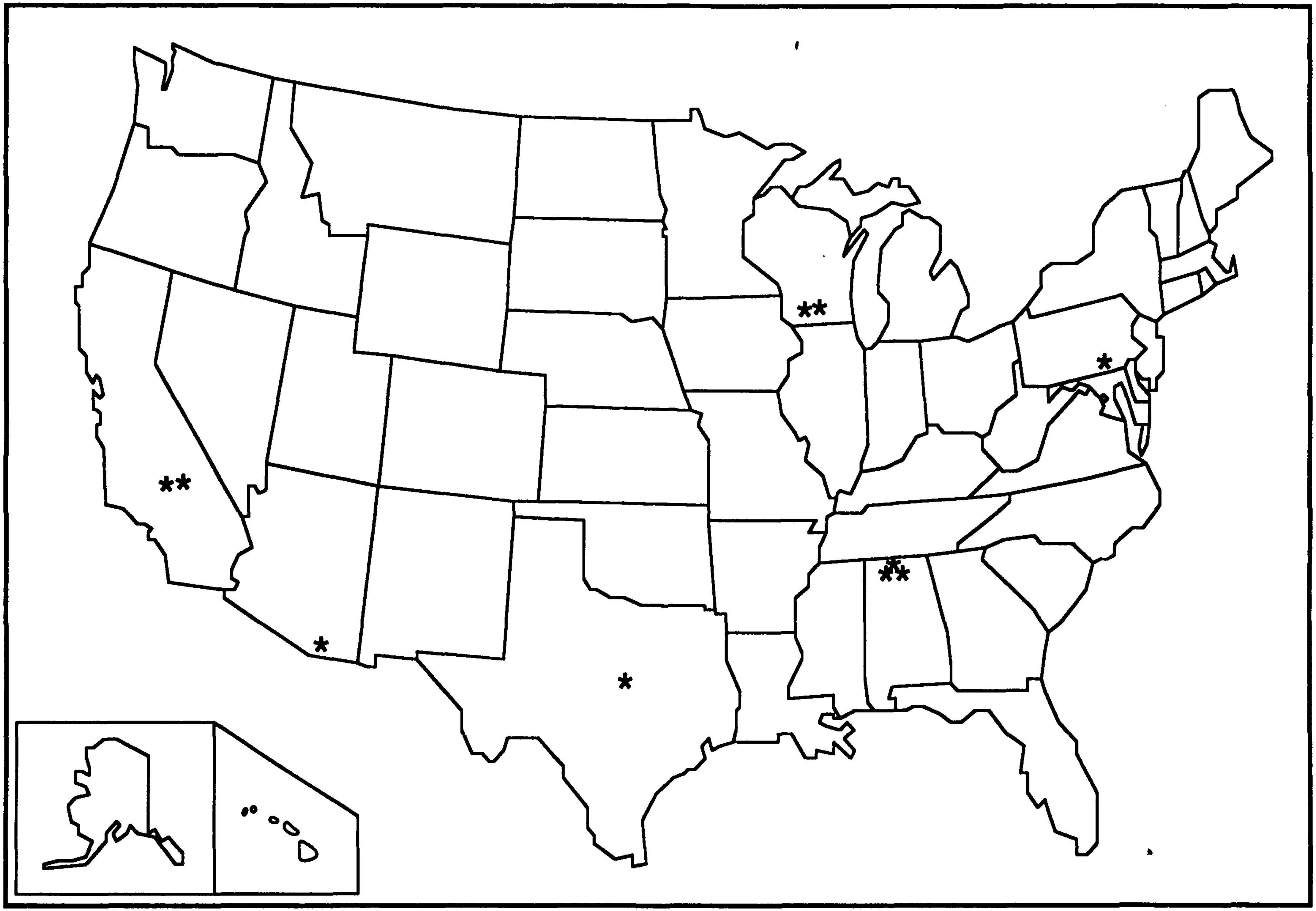
Prototype Categorization System

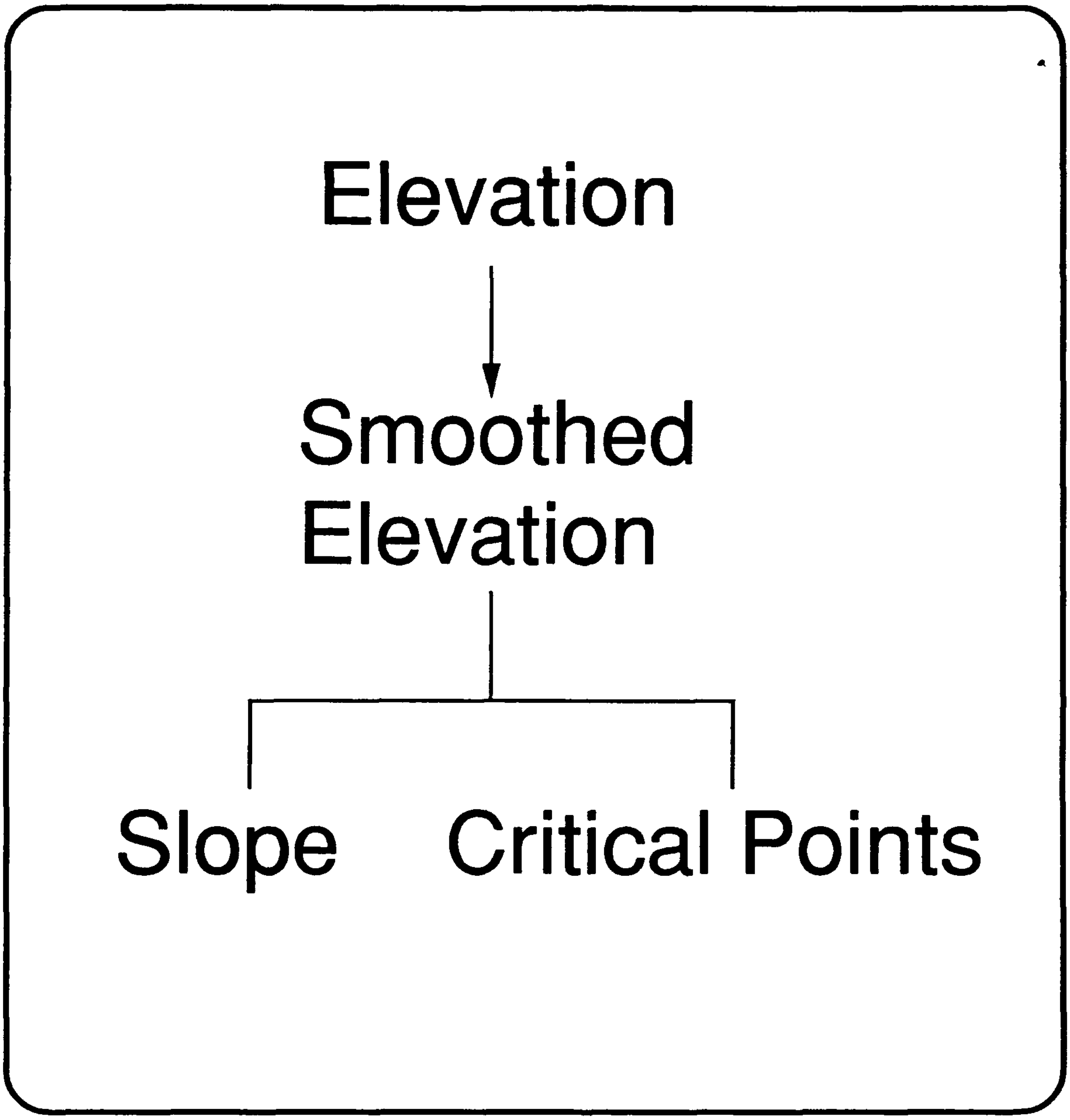


Categorization System for Terrain Features

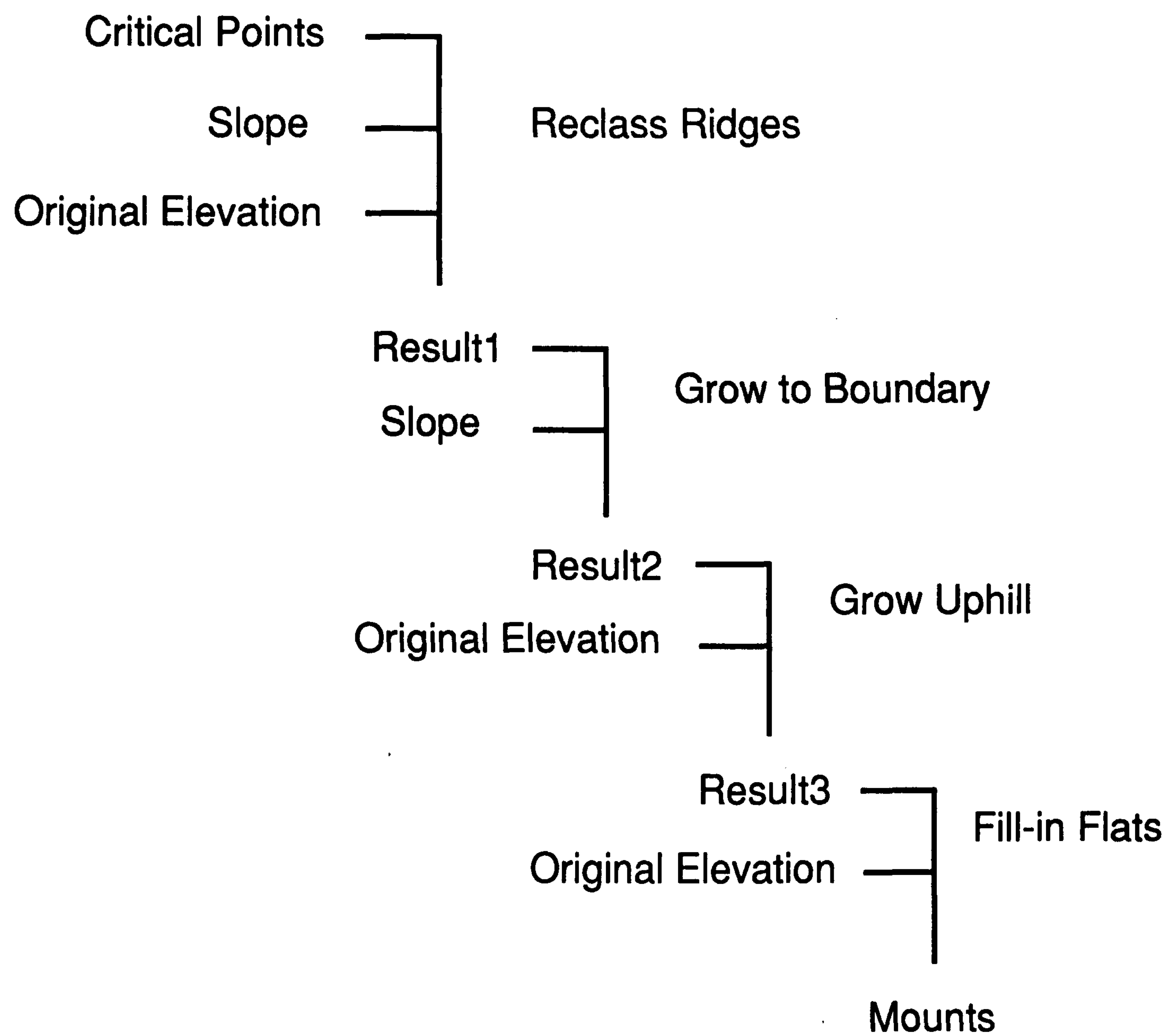


4





Preprocessing



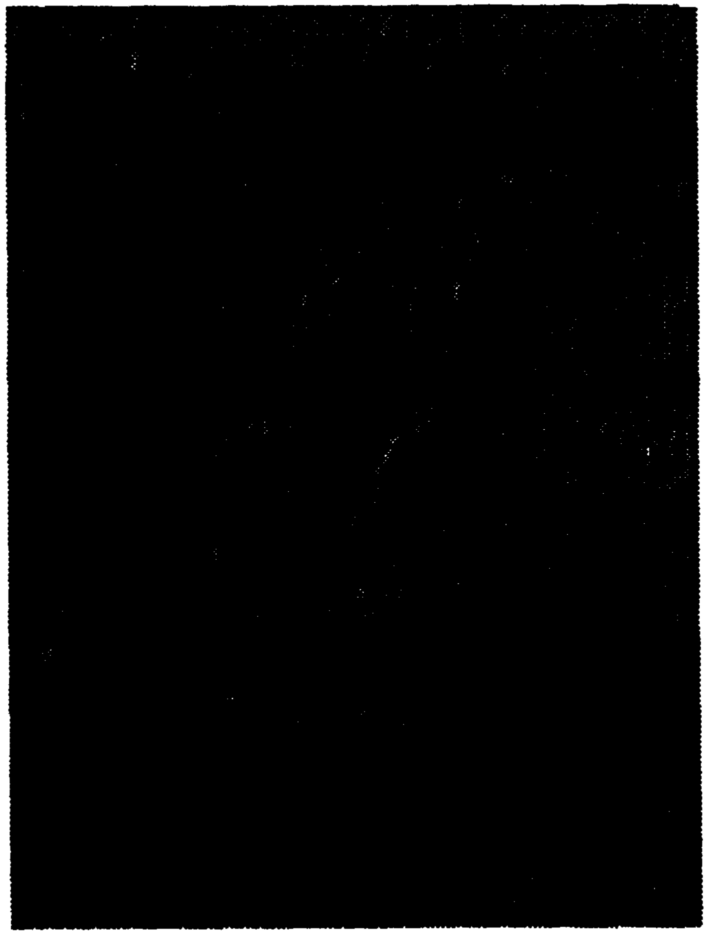
Critical Points
Slope
Original Elevation
Reclass Ridges
Result1
Slope
Grow to Boundary
Result2
Original Elevation
Grow Uphill
Result3
Original Elevation
Fill-in Flats
Mounts

Remove Clumps

Mounts $\longrightarrow$ Sieved Mounts

Postprocessing

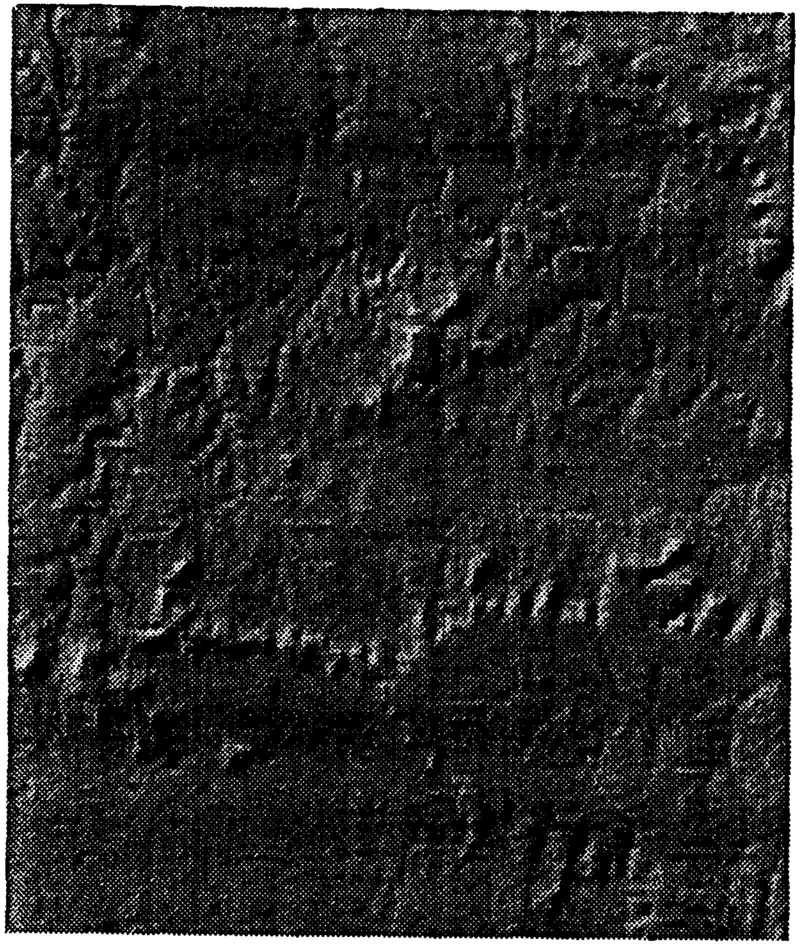
Gettysburg, PA



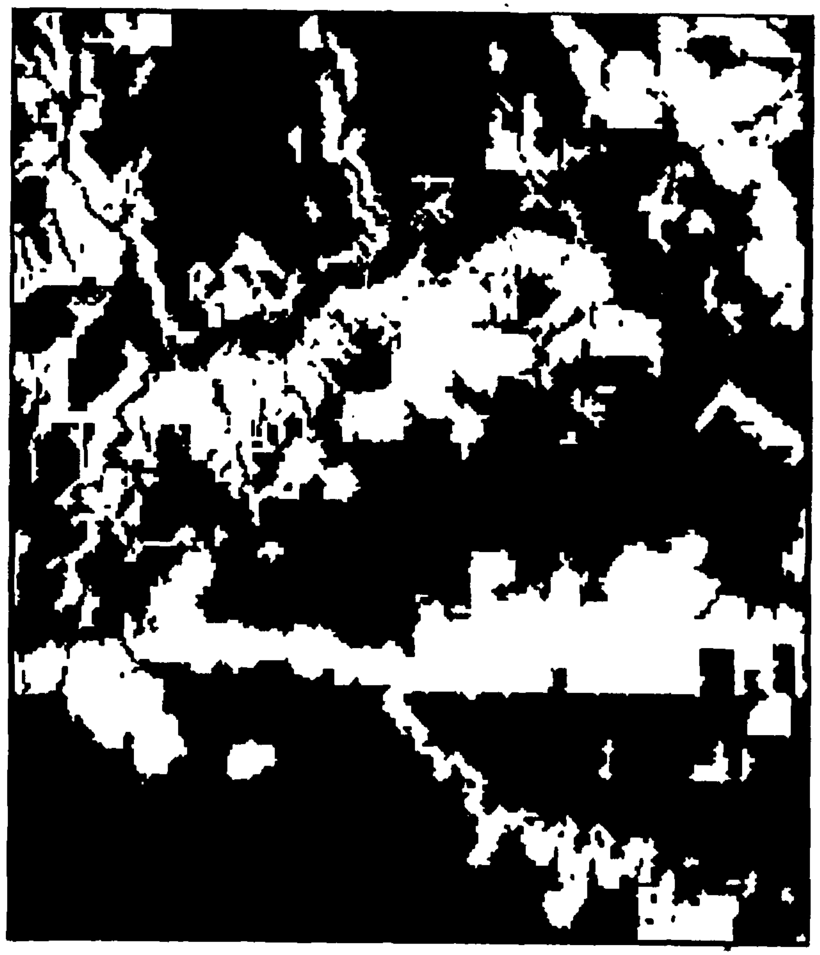
Shaded Relief

Classified Mounts

Post Oak Mountains, TX



Shaded Relief



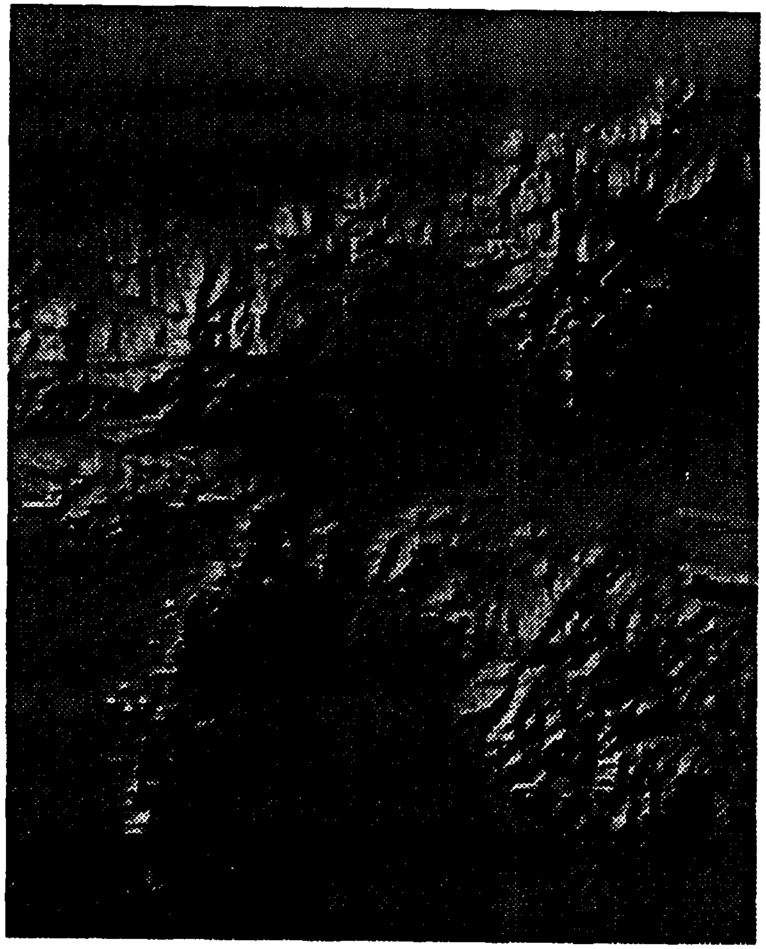
Classified Mounts

Mustang Mountains, AZ

Shaded Relief

Classified Mounts

Drinkwater Lake, CA



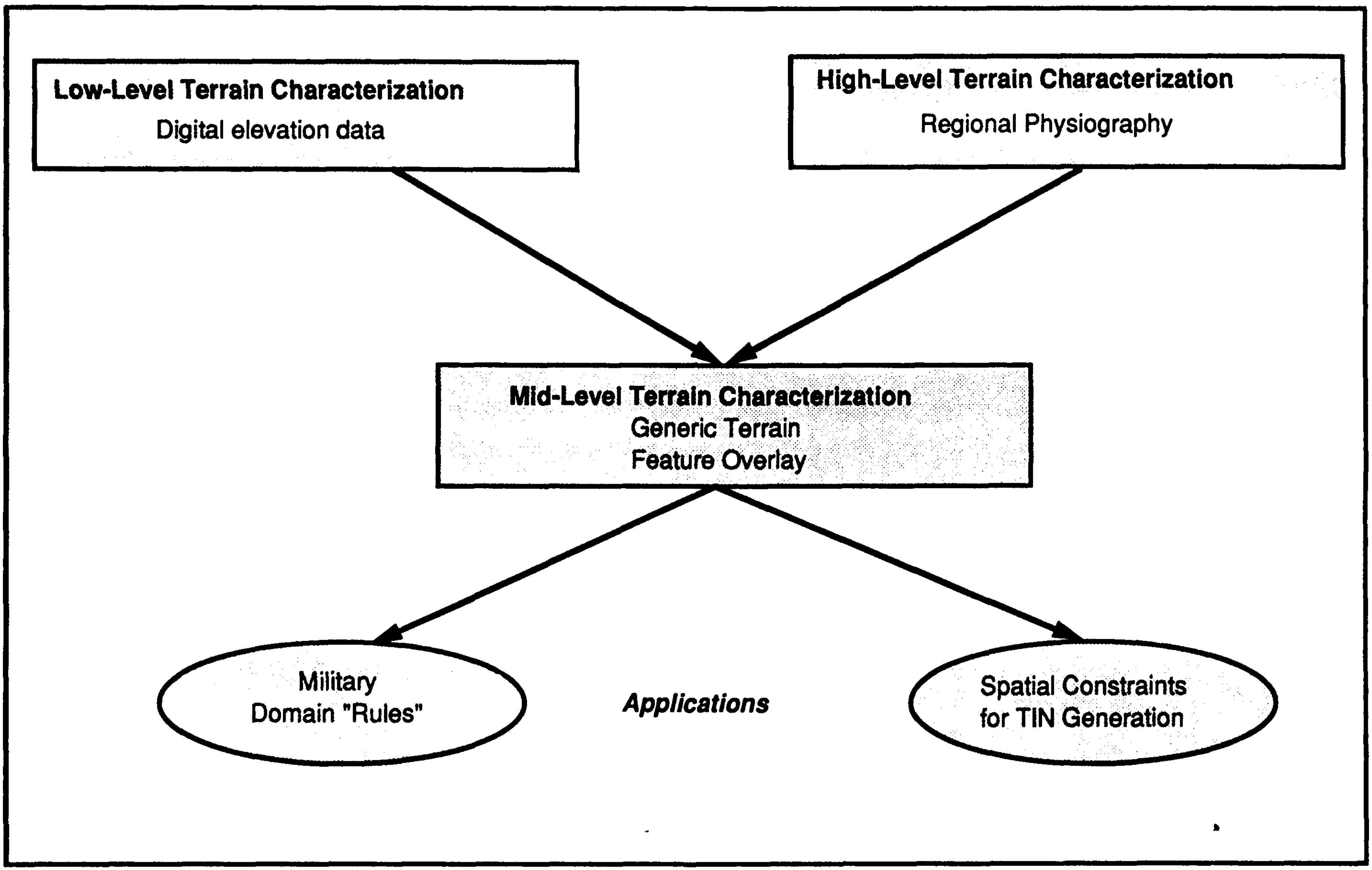
Shaded Relief

Classified Mounts

H

Method Limitations

- *Universal Approach*
- *Local Neighborhood Operators (3x3)*
- *"Critical Values"*


Low-Level Terrain Characterization
Digital elevation data
High-Level Terrain Characterization
Regional Physiography
Mid-Level Terrain Characterization
Generic Terrain
Feature Overlay
Military
Domain "Rules"
Applications
Spatial Constraints
for TIN Generation